AD-A104 607 PITTSBURGH UNIV PA DEPT OF MATHEMATICS AND STATISTICS F/G 12/1
SOME CONCEPTS OF NEGATIVE DEPENDENCE.(U)
MAR 80 H W BLOCK, T H SAVITS, M SHAKED N00014-76-C-0839

UNCLASSIFIED NL

1 of 1
AD A
104607

END
DATE
FILMED
10-81
DTIC

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

LEVEL II

12

AD A104607

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A104607	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Some Concepts of Negative Dependence		5. TYPE OF REPORT & PERIOD COVERED Technical
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Henry W. Block, Thomas H. Savits, Moshe Shaked		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0839, NR 042-357
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of Pittsburgh, Department of Statistics, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 12 39
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research (411(SP)), 800 N. Quincy Street, Arlington, VA 22217		12. REPORT DATE March 1980
		13. NUMBER OF PAGES 36
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC ELECTE SEP 28 1981 E

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Multinomial	Multivariate Hypergeometric	Negative and Positive Dependence
Dirichlet	Multivariate Normal	New Better Than Used
Inequalities	PF Densities	
Stochastic Ordering	Total Positivity	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The theory of positive dependence notions cannot yield useful results for some widely used distributions such as the multinomial, Dirichlet and the multivariate hypergeometric. Some conditions of negative dependence that are satisfied by these distributions and which have practical meaning are introduced. Preservation results for some of these concepts are derived. Useful inequalities for some widely used distributions are obtained. Results of Mallows (1969) that apply to the multinomial distributions are extended to more distributions. Examples are listed.

DD FORM 1473 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

DTIC FILE COPY

410268

8

12

SOME CONCEPTS OF NEGATIVE
DEPENDENCE

by

Henry W. Block[1,3]

Thomas H. Savits[1,4]

and

Moshe Shaked[2]

March, 1980

Accession For		
NTIS GRA&I	X	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A		

[1]Department of Mathematics and Statistics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260.

[2]Department of Mathematics, Indiana University, Bloomington, Indiana 47505.

[3]Supported by ONR Contract N00014-76-C-0839.

[4]Supported by NSF Grant MCS77-01458 and ONR Contract N00014-76-C-0839.

81 9 28 005

Abstract

The theory of positive dependence notions cannot yield useful results for some widely used distributions such as the multinomial, Dirichlet and the multivariate hypergeometric. Some conditions of negative dependence that are satisfied by these distributions and which have practical meaning are introduced. Preservation results for some of these concepts are derived. Useful inequalities for some widely used distributions are obtained. Results of Mallows (1969) that apply to the multinomial distributions are extended to more distributions. Examples are listed.

AMS Subject Classification: Primary 62H20, 62H05, 62E10, Secondary 60E05.

Key words: Multinomial, Dirichlet, multivariate hypergeometric, multivariate normal, PF_2 densities, total positivity, negative and positive dependence, new better than used, stochastic ordering, inequalities.

1. Introduction.

Concepts of positive dependence of sets of random variables (rv's) have received a lot of attention recently. Their study was found to yield a better understanding of the structure of some widely used multivariate distribution functions (df's). In addition to this, various useful inequalities were obtained with applications in many areas of probability and statistics. Barlow and Proschan (1975), Ch. 5, include a review of most of the work done prior to 1972. A list of more recent references can be found in Ahmed et al. (1978).

On the other hand notions of negative dependence have received very little attention in the literature. Some negative dependence analogs of positive dependence concepts have been mentioned by some authors (Lehmann (1966), Brindley and Thompson (1972), Dykstra, Hewett and Thompson (1973) and Shaked (1977) among others). In the bivariate setting the random vector (T_1, T_2) is usually said to satisfy some negative dependence condition if $(T_1, -T_2)$ satisfies the analogous positive dependence condition. However, this method of formulation cannot apply to higher dimensions. To the best of our knowledge Lehmann (1966) in the bivariate setting and Mallows (1968) in the multivariate setting came the nearest to a systematic study of negative dependence concepts; Mallows' discussion, however, is restricted to the multinomial distribution. The development below is in the spirit of Mallows; his results are special cases of ours.

While the first draft of this paper was being written two related works were brought to our attention. The first work by Ebrahimi and Ghosh (1980) discusses some negative dependence analogs of well known positive dependence concepts. Some of our definitions overlap those of Ebrahimi and Ghosh (1980); however, our main results differ from theirs. The second related paper is by Karlin and Rinott (1980). They introduce a negative dependence notion which is closely related to one of ours and they obtain some results which are similar to ours. Some remarks about the relationship between these two works and the present paper will be given throughout the text.

The main motivation for our definitions is to try to formulate the intuitive requirement that if a set of negatively dependent random variables is split into two subsets in some manner then one subset will tend to be 'large' when the other subset is 'small' and vice versa. In Section 2 we define the conditions to be discussed. We derive some inequalities in Section 3 and prove some preservation properties in Section 4. Examples are given in Section 5.

In the following "increasing" stands for "nondecreasing" and "decreasing" for "nonincreasing". Vectors in $\mathbb{R}^n$ are denoted by $\underline{t} = (t_1,\ldots,t_n)$ and $\underline{t} \leq \underline{t}'$ means $t_i \leq t_i'$, $i = 1,\ldots,n$. Similarly $\underline{t} < \underline{t}'$ means $t_i < t_i'$, $i = 1,\ldots,n$, and $\underline{0} = (0,\ldots,0)$. A real function on $\mathbb{R}^n$ will be called increasing if it is increasing in each variable when the other variables are held fixed.

A rv X is said to be stochastically smaller than the rv Y (denoted by $X \overset{st}{\leq} Y$) if $P(X > x) \leq P(Y > x)$ for every real x. The random vector $\underline{X} = (X_1,\ldots,X_n)$ is said to be stochastically smaller than $\underline{Y} = (Y_1,\ldots,Y_n)$ [denoted by $\underline{X} \overset{st}{\leq} \underline{Y}$] if $g(\underline{X}) \overset{st}{\leq} g(\underline{Y})$ for every $g \in C$ where C is the class of Borel measurable increasing functions on $\mathbb{R}^n$. If $\underline{X}$ and $\underline{Y}$ have the same df then we write $\underline{X} \overset{st}{=} \underline{Y}$. It is well known that $\underline{X} \overset{st}{\leq} \underline{Y}$ if and only if

$$(1.1) \qquad P(\underline{X} \in U) \leq P(\underline{Y} \in U) \quad \text{for every upper Borel set } U \text{ in } \mathbb{R}^n.$$

[U is an upper set if $\underline{x} \in U$ and $\underline{x} \leq \underline{y}$ implies that $\underline{y} \in U$.] According to Kamae, Krengel and O'Brien (1977), we need only consider open upper sets U in $\mathbb{R}^n$. It is also well known that for every random vector $\underline{X}$,

(1.2) $$\underline{X} + \underline{a} \overset{st}{\geq} \underline{X} \quad \text{whenever} \quad \underline{a} \geq \underline{0}$$

and that

(1.3) $$\underline{X} + \underline{A} \overset{st}{\geq} \underline{X} \quad \text{whenever } \underline{A} \text{ is a nonnegative random vector.}$$

Also if $P(\underline{X} \geq \underline{0}) = 1$ then

(1.4) $$a\underline{X} \overset{st}{\geq} \underline{X} \quad \text{whenever} \quad a \geq 1.$$

See Arjas and Lehtonen (1978) for an excellent review on stochastic ordering.

If $\underline{X}$ and $\underline{Y}$ are random vectors such that $\underline{X}$ given that $\underline{Y} = \underline{y}$ is stochastically smaller (larger) than $\underline{X}$ given that $\underline{Y} = \underline{y}'$, whenever $\underline{y} \leq \underline{y}'$ and $\underline{y}$ and $\underline{y}'$ are in the support of $\underline{Y}$, then we write

$$[\underline{X} | \underline{Y} = \underline{y}] \overset{st}{\uparrow} (\overset{st}{\downarrow})\ \underline{y}.$$

More precisely, this means that for every upper set U, there exists a version of $P(\underline{X} \in U | \underline{Y}) = \phi(\underline{Y})$ such that $\phi(\underline{y})$ is increasing (decreasing) in $\underline{y}$ on the support of $\underline{Y}$.

A bivariate function $K(\cdot,\cdot)$ which is defined on $S_1 \times S_2$ (where S_1 and S_2 are subsets of $\mathbb{R}$) is said to be <u>totally positive of order 2</u> (TP_2) on $S_1 \times S_2$ if $K(x,y) \geq 0$ and if

(1.5) $$K(x,y)\, K(x',y') \geq K(x,y')\, K(x',y) \quad \text{whenever} \quad x \leq x', \quad y \leq y'$$

(see Karlin (1968)). The function K is said to be reverse regular of order 2 (RR_2) on $S_1 \times S_2$ if $K(x,y) \geq 0$ and if

(1.6) $$K(x,y)\, K(x',y') \leq K(x,y')\, K(x',y) \quad \text{whenever} \quad x \leq x', \quad y \leq y',$$

(see Karlin (1968), p. 12).

2. Negative dependence concepts.

Most, but not all of the positive dependence concepts discussed in the literature have negative dependence analogs that can be obtained by changing the direction of the monotonicity or of the inequalities which define them. Here we define four conditions of this type which we find useful because we have methods of identifying distributions which satisfy them. Two of them are direct analogs of "conditionally increasing in sequence" and "positive orthant dependent". The third has a positive dependent counterpart, but it does not seem to have been discussed in the literature, while the fourth is a variation of "totally positive of order two (TP_2) in pairs".

In the case of positive dependence, one of the strongest and most useful notions is that of TP_2-ness in pairs; that is, the joint density or the discrete probability function f is assumed to exist and be TP_2 in pairs (see Barlow and Proschan (1975), p. 149). The natural negative dependence analog then is to assume that f is RR_2 in pairs. There are several drawbacks to this notion, however. Firstly, unlike the situation in which the joint density (or discrete probability function) is TP_2 in pairs, the marginal densities do not necessarily enjoy the same property. A simple $3 \times 2 \times 2$ discrete example suffices to show this. In fact, Theorem 5.1, p. 123 of Karlin (1968) is false when TP_2 is replaced by RR_2. (It should also be remarked here that even in the TP_2 case, one must make some assumptions on the nature of the set $\{f > 0\}$ before one can use the above Theorem 5.1 to conclude that the marginal densities are also TP_2 in pairs: see Kemperman (1977)). A possible alternative then is to assume that not only f, but all its marginal densities are RR_2 in pairs. But even under this assumption we have not been able to show that our weakest condition (see Definition 2.3) is consequently satisfied. Ebrahimi and Ghosh (1980) claim to have proven this result; however, their proof is based on an implication which, as will be shown below, does not hold (see discussion after Definition 2.4).

Because of these drawbacks and since we do not always want to assume the existence of a density, we prefer to work directly with the measure itself. This point of view is consistent with our other definitions in that we do not make any assumptions on the existence of a density in defining them. We are thus led to our first definition.

Let μ be a probability measure on the Borel sets in $\mathbb{R}^n$. If $I_1,\ldots,I_n$ are intervals in $\mathbb{R}^1$, we define the set function $\tilde{\mu}(I_1,\ldots,I_n)$ by $\tilde{\mu}(I_1,\ldots,I_n) = \mu(I_1\times\ldots\times I_n)$. By abuse of notation, we write μ instead of $\tilde{\mu}$. If I and J are intervals in $\mathbb{R}^1$, we write $I < J$ if $x \in I$, $y \in J$ implies $x < y$, that is, I lies to the left of J.

Definition 2.1. Let μ be a probability measure on $\mathbb{R}^2$. We say that μ is reverse regular of order two (RR_2) if

$$(2.1) \qquad \mu(I_1,I_2)\,\mu(I_1',I_2') \leq \mu(I_1,I_2')\,\mu(I_1',I_2)$$

for all intervals $I_1 < I_1'$, $I_2 < I_2'$ in $\mathbb{R}^1$. We also say that $\mu(I_1,I_2)$ is RR_2 in the variables I_1,I_2. If μ is a probability measure on $\mathbb{R}^n$

($n \geq 2$), we say that μ is <u>RR_2 in pairs</u> if $\mu(I_1,\ldots,I_n)$ is RR_2 in the pairs I_i, I_j for all $1 \leq i < j \leq n$ when the remaining variables are held fixed. The random variables $T_1,\ldots,T_n$ (or the random vector $\underline{T}$ or its distribution function F) are said to be <u>RR_2 in pairs</u> if its corresponding probability measure on $\mathbb{R}^n$ is.

<u>Remarks</u>:

(i) The obvious TP_2 definitions for μ are obtained by reversing the inequality in (2.1).

(ii) Clearly, if μ is $RR_2(TP_2)$ in pairs, then so are all marginals. Furthermore, it is not difficult to show that if F is the distribution function associated with μ and if $\overline{F}(t_1,\ldots,t_n) = \mu((t_1,\infty), (t_2,\infty),\ldots,(t_n,\infty))$ is the survival function, then the functions F and $\overline{F}$ are $RR_2(TP_2)$ in pairs in the sense of (1.5) and (1.6).

(iii) It is easy to show by a simple limiting argument that if μ is $RR_2(TP_2)$ in pairs, and if μ has a density f with respect to a product measure $m = m_1\times\ldots\times m_n$ of σ-finite measures such that f is continuous on the support of m and zero off the support of m, then f is $RR_2(TP_2)$ in pairs.

(iv) In the $n = 2$ case, we have the stronger converse, namely, if μ has a density f with respect to a product measure $m = m_1 \times m_2$ of σ-finite measures which is $RR_2(TP_2)$ on $S_1 \times S_2$, where S_i is the support of $m_i (i = 1,2)$, then μ is $RR_2(TP_2)$.

(v) In the TP_2 case, one can generalize to higher dimensions if one makes some assumption on the set $\{f > 0\}$. Let μ have a density f with respect to a product measure $m = m_1\times\ldots\times m_n$ of

σ-finite measures. Let S_i be the support of m_i. Then the support of m is $S = S_1 \times \ldots \times S_n$. We assume that there exists $\tilde{S} = \tilde{S}_1 \times \ldots \times \tilde{S}_n$ such that $\{f > 0\} \cap S = \tilde{S}$ and that f is TP_2 in pairs on $\tilde{S}$. Then μ is TP_2 in pairs. Just use Theorem 5.1, page 123 of Karlin (1968) repeatedly to show, e.g., that for fixed intervals $I_3, \ldots, I_n$ in $\mathbb{R}^1$,

$$g(x_1, x_2) = \int_{I_3} \ldots \int_{I_n} f(x_1, x_2, x_3, \ldots, x_n)\, dm_3(x_3) \ldots dm_n(x_n)$$

is TP_2 in x_1 and x_2 on $\tilde{S}_1 \times \tilde{S}_2$. The result then follows by a simple integration using the TP_2 inequality for g.

(vi) The generalization to the RR_2 case is not as simple. If one assumes, however, that μ has a density f with respect to a product measure $m = m_1 \times \ldots \times m_n$ of σ-finite measures such that the density f when integrated over any $n-2$ intervals in $\mathbb{R}^1$ is RR_2 in the remaining unintegrated variables, then μ is RR_2 in pairs. In terms of random variables, this can be paraphrased as follows. Let $T_1, \ldots, T_n$ be random variables with a density f (with respect to a product measure of σ-finite measures). Then μ is RR_2 in pairs if and only if for every $1 \leq i < j \leq n$ the conditional density of

$$(T_i, T_j) \Big| \bigcap_{k \neq i,j} \{T_k \in I_k\}$$

is RR_2 in t_i and t_j for all choices of intervals $I_k (k \neq i,j)$ in $\mathbb{R}^1$. Equivalently, if χ_I denotes the indicator function of I, then μ is RR_2 in pairs if and only if

$$\int_{\mathbb{R}} \ldots \int_{\mathbb{R}} \Big[\prod_{k \neq i,j} \chi_{I_k}(t_k) \Big] f(t_1, \ldots, t_n) \Big[\prod_{k \neq i,j} dt_k \Big]$$

is RR_2 in the unintegrated variables t_i and t_j for all choices of intervals I_k $(k \neq i,j)$ in $\mathbb{R}^1$. By replacing χ_{I_k} by ϕ_k in the above integral, and requiring it to be RR_2 in t_i and t_j whenever $\{\phi_k\}_{k \neq i,j}$ is a set of PF_2 functions, one obtains the negative dependence condition of Karlin and Rinott (1980). It is a condition which is stronger than the one of Definition 2.1 as can be easily seen by recalling that the indicator function of an interval is PF_2.

(vii) Clearly, if $\{\mu_n\}$ is a sequence of RR_2 in pairs probability measures and if μ_n converges weakly to μ then μ is RR_2 in pairs.

Definition 2.2. The rv's $T_1,\ldots,T_n$ (or the random vector $\underline{T}$ or its df) are said to be conditionally decreasing in sequence (CDS) if, for $i = 1,2,\ldots,n-1$,

(2.2) $$[T_{i+1} | T_1 = t_1,\ldots,T_i = t_i] \overset{st}{\downarrow} (t_1,\ldots,t_i).$$

Definition 2.3. The rv's $T_1,\ldots,T_n$ (or the random vector $\underline{T}$ or its df) are said to be negatively upper orthant dependent (NUOD) if for every $\underline{t}$,

(2.3) $$P(\underline{T} > \underline{t}) \leq \prod_{i=1}^{n} P(T_i > t_i).$$

They are said to be negatively lower orthant dependent (NLOD) if for every $\underline{t}$,

(2.3') $$P(\underline{T} \leq \underline{t}) \leq \prod_{i=1}^{n} P(T_i \leq t_i).$$

When $n = 2$, (2.3) and (2.3') are equivalent, but not when $n \geq 3$ (see, for example, Ebrahimi and Ghosh (1980)).

The next concept has a natural positive dependence analog, however, we are not aware of any place in the literature in which it has been discussed.

Definition 2.4. The rv's $T_1,\ldots,T_n$ (or the random vector $\underline{T}$ or its df) are said to be negatively dependent in sequence (NDS) if, for $i = 2,3,\ldots,n$,

$$(2.4) \qquad [(T_1,\ldots,T_{i-1})|T_i = t_i] \overset{st}{\downarrow} t_i.$$

Often, to verify (2.4), one can find it easier to verify that for $i = 1,\ldots,n$,

$$(2.4') \qquad [(T_1,\ldots,T_{i-1},T_{i+1},\ldots,T_n)|T_i = t_i] \overset{st}{\downarrow} t_i.$$

Then clearly (2.4) holds.

We now investigate some of the relationships among the various definitions. First note that NDS implies both NUOD and NLOD and these implications are sharp. To see this, use methods similar to Barlow and Proschan (1975) to show that (2.4) implies for $i = 1,\ldots,n-1$,

$$(2.5) \qquad [(T_1,\ldots,T_i)|T_{i+1} > t_{i+1}] \overset{st}{\geq} [(T_1,\ldots,T_i)|T_{i+1} > t'_{i+1}]$$

whenever $t_{i+1} \leq t'_{i+1}$. (Although Barlow and Proschan assumed the existence of a density, a modification of their proof works.) But from (2.5) it follows that

$$P(T_1 > t_1, \ldots, T_n > t_n) \le P(T_1 > t_1, \ldots, T_{n-1} > t_{n-1})\, P(T_n > t_n)$$

$$\le P(T_1 > t_1, \ldots, T_{n-2} > t_{n-2}) \prod_{i=n-1}^{n} P(T_i > t_i)$$

$$\le \cdots \le \prod_{i=1}^{n} P(T_i > t_i)$$

which proves (2.3). The proof of (2.4) $\Rightarrow$ (2.3') is similar.

It is not difficult to construct an example which shows that CDS $\not\Rightarrow$ NUOD and that CDS $\not\Rightarrow$ NLOD and hence CDS $\not\Rightarrow$ NDS. For example, let $P(T_1=1, T_2=1) = P(T_1=1, T_2=2) = P(T_1=2, T_2=2) = .1$ and $P(T_1=2, T_2=1) = .7$ and let T_3 given $T_1 = 1$, $T_2 = 1$ be degenerate at 11, T_3 given $T_1 = 1$, $T_2 = 2$ be degenerate at 1, T_3 given $T_1 = 2$, $T_2 = 1$ be degenerate at 10 and T_3 given $T_1 = 2$, $T_2 = 2$ be degenerate at 1. Then, clearly, (T_1, T_2, T_3) is CDS but $P(T_1 > 1, T_3 > 1) = .7 > .64 = P(T_1 > 1)\, P(T_3 > 1)$, thus (T_1, T_2, T_3) is neither NUOD nor NLOD. Ebrahimi and Ghosh (1980) claim that CDS $\Rightarrow$ NUOD; the example shows that this is not the case.

Next we show that NDS $\not\Rightarrow$ CDS. Let (T_1, T_2, T_3) take on values on the eight vertices of the unit cube such that $P(T_1=1, T_2=0, T_3=1) = P(T_1=1, T_2=1, T_3=0) = .2$ and the other six probabilities are .1. It is easy to see that T_1 and T_2 are independent and that $[(T_1, T_2) | T_3 = 0] \overset{st}{\ge} [(T_1, T_2) | T_3 = 1]$. Thus, (T_1, T_2, T_3) is NDS. But $P(T_3 > 0 | T_1 = 0, T_2 = 0) = 1/2 < 2/3 = P(T_3 > 0 | T_1 = 1, T_2 = 0)$, hence (T_1, T_2, T_3) is not CDS.

We will shortly show that under some reasonable assumption, $RR_2 \Rightarrow$ CDS. It is not known whether $RR_2 \Rightarrow$ NDS, but it does imply NUOD and NLOD. This is clear since from the remark (ii) following Definition 2.1, we have that both F and $\bar{F}$ are RR_2 in pairs. The result then follows by a simple argument.

The implications $RR_2 \Rightarrow CDS$, $RR_2 \Rightarrow NUOD$, $RR_2 \Rightarrow NLOD$, $NDS \Rightarrow NUOD$ and $NDS \Rightarrow NLOD$ justify the consideration of the RR_2 and the NDS concepts. In Sections 4 and 5 it will be shown that many df's are RR_2 or NDS and thus these df's satisfy the meaningful CDS concept and the inequalities which can be derived from the NUOD and NLOD concepts. In addition to it the NDS concept is intuitively meaningful by itself. We mention, in passing, that since Karlin and Rinott's (1980) condition implies the RR_2 condition it follows that it implies the NUOD and the NLOD inequalities (2.3) and (2.3'). Actually, Karlin and Rinott (1980) obtained some additional useful inequalities which follow from their stronger condition.

To justify calling (2.1)-(2.5) "conditions for negative dependence" we have to show that they imply

$$\text{(2.6)} \qquad COV(T_i, T_j) \le 0, \quad 1 \le i < j \le n,$$

when the second moments exist.

From (2.3), it follows that $P(T_i > t_i, T_j > t_j) \le P(T_i > t_i)\, P(T_j > t_j)$ and it is well known that this inequality implies (2.6) [see e.g., Lehmann (1966)]. Similarly (2.3') implies (2.6) and hence also (2.4) and (2.5) imply (2.6).

If $(T_1, \ldots, T_n)$ is CDS then $[T_2 | T_1 = t_1] \overset{st}{\downarrow} t_1$; hence $cov(T_1, T_2) \le 0$. Thus, if $(T_{\pi(1)}, \ldots, T_{\pi(n)})$ is CDS for every permutation π of $\{1,2,\ldots,n\}$ then (2.6) holds.

We close this section with a proof that RR_2 in pairs implies CDS under some reasonable assumptions. Let μ be RR_2 in pairs and let $(T_1, \ldots, T_n)$

be any random vector having μ as its induced probability measure. Fix $t_{i+1} \in \mathbb{R}$ and an integer i, $1 \le i \le n-1$. Define the set functions ν and λ on the Borel subsets of $\mathbb{R}^i$ by

$$\nu(A) = P(T_{i+1} > t_{i+1};\ (T_1,\ldots,T_i) \in A)$$

$$\lambda(A) = P((T_1,\ldots,T_i) \in A).$$

Note that λ is just a marginal of μ. Then if $\{R_{ik}^{\ell}\}_{\ell=1}^{\infty}$ is a sequence of rectangles which partition $\mathbb{R}^i$ for each ℓ and whose mesh size tends to zero as $\ell \to \infty$, and if $\{R_{ik}^{\ell+1}\}$ is a refinement of $\{R_{ik}^{\ell}\}$ for all ℓ, then it is well known by martingale arguments (see, e.g., Meyer (1966)) that on a set D with $\lambda(D^c) = 0$, $\nu(R_{ik}^{\ell})/\lambda(R_{ik}^{\ell}) \to \phi(t_1,\ldots,t_i)$ as R_{ik}^{ℓ} decreases to $(t_1,\ldots,t_i)$ pointwise and in L^1; moreover $\phi(t_1,\ldots,t_i)$ is a version of $P(T_{i+1} > t_{i+1} | T_1 = t_1,\ldots,T_i = t_i)$. We may assume without loss of generality that $D \subseteq$ support λ. We want to show that ϕ is decreasing on D. This will follow if we can assume that the support Λ of λ satisfies a chain condition; that is, if $\underline{t} \in \Lambda$ and $\underline{t}' \in \Lambda$ with $\underline{t} \le \underline{t}'$, then there exist $\underline{t}_j \in \Lambda$ such that $\underline{t} = \underline{t}_0 \le \underline{t}_1 \le \ldots \le \underline{t}_m = \underline{t}'$ and $\underline{t}_j$ differs from $\underline{t}_{j+1}$ in only one component. To see this, suppose that $\underline{t}_0$ and $\underline{t}_1$ differ in only the first component. Then let $I_1 < I_1'$, $I_{i+1} = (-\infty, t_{i+1}] < I_{i+1}' = (t_{i+1}, \infty)$ and $I_2,\ldots,I_i$ be any intervals such that $\underline{t}_0 \in I_1 \times I_2 \times \ldots \times I_i$, $\underline{t}_1 \in I_1' \times I_2 \times \ldots \times I_i$. If ρ is the marginal of $(T_1,\ldots,T_{i+1})$, then since it is RR_2 in pairs, we have

$$0 \le \begin{vmatrix} \rho(I_1,I_2,\ldots,I_{i+1}') & \rho(I_1',I_2,\ldots,I_{i+1}') \\ \rho(I_1,I_2,\ldots,I_{i+1}) & \rho(I_1',I_2,\ldots,I_{i+1}) \end{vmatrix} = \begin{vmatrix} \rho(I_1,I_2,\ldots,I_{i+1}') & \rho(I_1',I_2,\ldots,I_{i+1}') \\ \rho(I_1,I_2,\ldots,\mathbb{R}^1) & \rho(I_1',I_2,\ldots,\mathbb{R}^1) \end{vmatrix}$$

or that

$$\frac{\nu(I_1,I_2,\ldots,I_i)}{\lambda(I_1,I_2,\ldots,I_i)} \geq \frac{\nu(I_1',I_2,\ldots,I_i)}{\lambda(I_1',I_2,\ldots,I_i)} .$$

Note that there is no trouble dividing since the denominators are non zero. Iterating this procedure to pass from $\underline{t}_1$ to $\underline{t}_2$, $\underline{t}_2$ to $\underline{t}_3,\ldots,\underline{t}_{m-1}$ to $\underline{t}_m$ and then letting the intervals shrink we find that $\phi(\underline{t}) \geq \phi(\underline{t}')$. Now set

$$\phi^*(\underline{t}) = \begin{cases} \phi(\underline{t}) & \text{if } \underline{t} \in D \\ \inf\{\phi(\underline{s});\ \underline{s} \in D,\ \underline{s} \leq \underline{t}\} \quad (\inf \emptyset = +\infty) & \text{if } \underline{t} \notin D. \end{cases}$$

It easily follows that ϕ^* is decreasing everywhere and is a version of $P(T_{i+1} > t_{i+1} | T_1 = t_1,\ldots,T_i = t_i)$.

The chain condition is easily seen to be satisfied if the support of μ is a cross product. Without some type of chain condition, we can only show that ϕ is decreasing componentwise on D. In this case, one may not be able to extend ϕ such that it is decreasing everywhere and is still a version of the condition probability.

3. Some inequalities.

This section is devoted to the derivation of some inequalities which may be of special interest. The results closely parallel those found in Karlin and Rinott (1980), but are derived under weaker assumptions.

Proposition 3.1. $(T_1,\ldots,T_n)$ is NUOD if and only if

$$E[\prod_{i=1}^{n} \phi_i(T_i)] \leq \prod_{i=1}^{n} E[\phi_i(T_i)] \tag{3.1}$$

whenever all ϕ_i are nonnegative and increasing. The result (3.1) is also true if we replace NUOD by NLOD and increasing by decreasing.

Proof. Let $\phi_i(t) = \chi_{(b_i,\infty)}(t)$, $i = 1,\ldots,n$. Then (3.1) reduces to the NUOD inequality (2.3). Since each side of (3.1) is multilinear in the ϕ_i, the result holds for nonnegative linear combinations of such indicator functions and hence for the general ϕ_i by a standard limiting argument. ||

Now suppose that μ is RR_2 in pairs. For every i, let $I_i = J_i \cup K_i$, all intervals, with $J_i < K_i$.

Theorem 3.1. If $1 \le k \le n$, then

$$(3.2) \qquad \mu(J_1,\ldots,J_n)\,\mu(I_1,\ldots,I_n) \le \mu(J_1,\ldots,J_k,I_{k+1},\ldots,I_n)$$

$$\times\, \mu(I_1,\ldots,I_k,J_{k+1},\ldots,J_n).$$

The result (3.2) is also true if we replace all J's by K's.

Proof. We proceed by induction. If $n = 2$, then

$$\mu(J_1,J_2)\,\mu(I_1,I_2) \le \begin{cases} \mu(J_1,I_2)\,\mu(I_1,J_2) & \text{if } k = 1. \\ \mu(J_1,J_2)\,\mu(I_1,I_2) & \text{if } k = 2. \end{cases}$$

The case $k = 1$ follows from the RR_2 assumption since

$$0 \ge \begin{vmatrix} \mu(J_1,J_2) & \mu(J_1,K_2) \\ \mu(K_1,J_2) & \mu(K_1,K_2) \end{vmatrix} = \begin{vmatrix} \mu(J_1,J_2) & (J_1,J_2\cup K_2) \\ (J_1\cup K_1,J_2) & (J_1\cup K_1,J_2\cup K_2) \end{vmatrix},$$

and the case $k = 2$ is an identity.

Now suppose that (3.2) is true whenever ν is a probability measure on $\mathbb{R}^n$ which is RR_2 in pairs. Let μ be a probability measure on $\mathbb{R}^{n+1}$ which is RR_2 in pairs and let $1 \leq k \leq n+1$. Since there is nothing to prove if $k = n + 1$, we may assume that $1 \leq k \leq n$. Similary, we may assume that $\mu(J_1,\ldots,J_n,J_{n+1}) \neq 0$. It suffices then to prove that

$$(3.3) \qquad \frac{\mu(I_1,\ldots,I_n,I_{n+1})}{\mu(J_1,\ldots,J_k,I_{k+1},\ldots,I_n,I_{n+1})} \leq \frac{\mu(I_1,\ldots,I_n,J_{n+1})}{\mu(J_1,\ldots,J_k,I_{k+1},\ldots,I_n,J_{n+1})}$$

since by the induction hypothesis, we have that

$$\frac{\mu(I_1,\ldots,I_n,J_{n+1})}{\mu(J_1,\ldots,J_k,I_{k+1},\ldots,I_n,J_{n+1})} \leq \frac{\mu(I_1,\ldots,I_k,J_{k+1},\ldots,J_n,J_{n+1})}{\mu(J_1,\ldots,J_n,J_{n+1})} .$$

But,

$$\frac{\mu(I_1,\ldots,I_n,I_{n+1})}{\mu(I_1,\ldots,I_n,J_{n+1})} \leq \frac{\mu(J_1,I_2,\ldots,I_n,I_{n+1})}{\mu(J_1,I_2,\ldots,I_n,J_{n+1})} \leq \cdots$$

$$\leq \frac{\mu(J_1,\ldots,J_k,I_{k+1},\ldots,I_n,I_{n+1})}{\mu(J_1,\ldots,J_k,I_{k+1},\ldots,I_n,J_{n+1})}$$

which is another way of writing (3.3). The jth inequality above follows from the fact that μ is RR_2 in the pair j and $(n+1)$. ||

Suppose that $(T_1,\ldots,T_n)$ is RR_2 in pairs.

<u>Corollary 1</u>. If α, β partition $\{1,\ldots,n\}$, then

$$(3.4)\qquad P(T_i\varepsilon J_i, i\varepsilon\alpha\cup\beta)\, P(T_i\varepsilon I_i, i\varepsilon\alpha\cup\beta)$$
$$\le P(T_i\varepsilon J_i, i\varepsilon\alpha; T_j\varepsilon I_j, j\varepsilon\beta)\, P(T_i\varepsilon I_i, i\varepsilon\alpha; T_j\varepsilon J_j, j\varepsilon\beta).$$

It also holds true if we replace all J's with the K's.

Corollary 2. If α, β, γ partition $\{1,\ldots,n\}$, then

$$(3.5)\qquad P(T_i\varepsilon L_i, i\varepsilon\alpha; T_j\varepsilon J_j, j\varepsilon\beta\cup\gamma)\, P(T_i\varepsilon L_i, i\varepsilon\alpha; T_j\varepsilon I_j, j\varepsilon\beta\cup\gamma)$$
$$\le P(T_i\varepsilon L_i, i\varepsilon\alpha; T_j\varepsilon J_j, j\varepsilon\beta; T_k\varepsilon I_k, k\varepsilon\gamma)\, P(T_i\varepsilon L_i, i\varepsilon\alpha; T_j\varepsilon I_j, j\varepsilon\beta; T_k\varepsilon J_k, k\varepsilon\gamma)$$

for any intervals L_i, $i \varepsilon \alpha$. It also holds true if we replace all J's with the K's.

If we take $J_i = (-\infty, b_i]$ and $I_i = (-\infty,\infty)$ for $i \varepsilon \alpha\cup\beta$ in Corollary 1, we get (3.5) of Karlin and Rinott (1980). If we take, in Corollary 2, $L_i = [a_i, b_i]$, $i \varepsilon \alpha$, $J_j = (-\infty, b_j]$ and $I_j = (-\infty,\infty)$ for $j \varepsilon \beta\cup\gamma$ we get (1.7) of Karlin and Rinott (1980).

Also, note that as soon as we have an inequality of the form

$$(3.6)\qquad P(T_1 \le b_1,\ldots,T_n \le b_n) \le P(T_1 \le b_1,\ldots,T_k \le b_k)\, P(T_{k+1} \le b_{k+1},\ldots,T_n \le b_n)$$

it follows as in the proof of Proposition 3.1 that

$$(3.7)\qquad E[\prod_{i=1}^{n} \phi_i(T_i)] \le E[\prod_{i=1}^{k} \phi_i(T_i)]\, E[\prod_{j=k+1}^{n} \phi_j(T_j)]$$

whenever ϕ_i are nonnegative and decreasing. Similarly, if we have

$$(3.8)\qquad P(T_1 > a_1,\ldots,T_n > a_n) \le P(T_1 > a_1,\ldots,T_k > a_k)\, P(T_{k+1} > a_{k+1},\ldots,T_n > a_n)$$

then (3.7) holds for ϕ_i nonnegative and increasing.

In particular, (3.6) holds if F is RR_2 in pairs and (3.8) holds if $\overline{F}$ is RR_2 in pairs. In fact, if F(or $\overline{F}$) is RR_2 in pairs, then it is MRR_2 in the sense of Karlin and Rinott (1980); i.e.,

$$F(\underline{x} \wedge \underline{y})\, F(\underline{x} \vee \underline{y}) \le F(\underline{x})\, F(\underline{y}).$$

4. Closure Results.

Preservation theorems are useful for identifying negatively dependent df's or for constructing new negatively dependent df's from known ones. In this section we discuss some preservation results and describe a method for the construction of negatively dependent df's.

Theorem 4.1. If $T_1,\ldots,T_n$ are (*) and if $\psi_1,\ldots,\psi_n$ are strictly increasing functions then $\psi_1(T_1),\ldots,\psi_n(T_n)$ are (*) where (*) is one of the following: RR_2 in pairs, CDS, NDS, NUOD or NLOD.

Theorem 4.2. If $(T_1,\ldots,T_n)$ and $(S_1,\ldots,S_n)$ are independent and are (*) then $(T_1,\ldots,T_n, S_1,\ldots,S_n)$ is (*) where (*) is the same as in Theorem 4.1.

The proofs of these theorems are straightforward and will be omitted.

The following preliminaries are needed for the statement of Theorem 4.3.

A univariate density f is said to be a Polya frequency function of order 2 (PF_2) if $f(x-y)$ is TP_2 on $\mathbb{R} \times \mathbb{R}$. A probability function f is PF_2 if $f(x-y)$ is TP_2 on $\mathbb{N} \times \mathbb{N}$ where $\mathbb{N} = \{\ldots,-1,0,1,\ldots\}$. A thorough discussion of PF_2 densities and many examples can be found in Karlin (1968).

It will be shown in a later section that many multivariate random variables satisfy a certain structural condition. In the next theorem we will show that this condition implies some of the dependence conditions that we have introduced.

Theorem 4.3. Let $S_0, S_1, \ldots, S_n$ be independent rv's and assume that each has a PF_2 density (or probability function). Fix s and let $(T_1, \ldots, T_n)$ have the same joint df as the conditional df of $(S_1, \ldots, S_n)$ given that $S_0 + S_1 + \ldots + S_n = s$, that is,

$$(4.1) \qquad (T_1, \ldots, T_n) \overset{st}{=} [(S_1, \ldots, S_n) | S_0 + S_1 + \ldots + S_n = s].$$

Then $(T_1, \ldots, T_n)$ is RR_2 in pairs and consequently CDS, NUOD and NLOD.

Proof. Let μ be the probability measure of $(T_1, \ldots, T_n)$ on $\mathbb{R}^n$. Then by assumption $\mu(I_1, \ldots, I_n) = P(S_1 \in I_1, \ldots, S_n \in I_n | S_0 + S_1 + \ldots + S_n = s)$. Now the joint density of $[(S_1, \ldots, S_n) | S_0 + S_1 + \ldots + S_n = s]$ is given by

$$c \prod_{i=1}^{n} f_i(s_i) f_0(s - s_1 - \ldots - s_n)$$

where c is a normalizing constant. We first show that μ is RR_2 in the variables I_1, I_2 when the remaining intervals $I_3, \ldots, I_n$ are held fixed. According to the remark (vi) following Definition 2.1, we need only show that

$$g(s_1, s_2) = c\, f_1(s_1)\, f_2(s_2)\, f_0'(s - s_1 - s_2)$$

is RR_2 in s_1 and s_2, where

$$f_0'(\zeta) = \int_{I_3} \cdots \int_{I_n} f_3(s_3) \cdots f_n(s_n)\, f_0(\zeta - s_3 - \cdots - s_n)\, dm(s_3) \cdots dm(s_n)$$

and m is either the Lebesgue measure or the counting measure. However, the above is nothing but the convolution of the PF_2 functions $f_3 \chi_{I_3}, \ldots, f_n \chi_{I_n}$ and f_0, where χ_A is the indicator function of the set A, and so f_0' is PF_2. It easily follows then that g is RR_2. The proof that μ is RR_2 in the variables I_i, I_j for all $1 \le i < j \le n$ is similar. ||

Remark. Actually, it is not difficult to show that under the assumptions of Theorem 4.3 the df of the random vector $(T_1, \ldots, T_n)$ satisfies the stronger condition of Karlin and Rinott (1980). The consequences of this observation are discussed below.

It is well known (see Section 5) that the multinomial, the multivariate hypergeometric, the Dirichlet and the Dirichlet compound multinomial random vectors as well as some multivariate normal random vectors with nonpositive correlations, can be represented as in (4.1). Thus, all the respective df's are RR_2 in pairs. Karlin and Rinott (1980) have shown that these df's actually satisfy their stronger condition; however, their proofs are quite involved and differ from one case to another. By the Remark after Theorem 4.3 these results of Karlin and Rinott (1980) follow at once.

The following theorem will be found useful in the next section.

Theorem 4.4. Assume that $(T_1, \ldots, T_n)$ and $(S_1, \ldots, S_n)$ are independent and NDS. If all the univariate marginal densities (with respect to Lebesgue measure) or probability functions in the discrete case of $\underline{S}$ and $\underline{T}$ are

PF_2, then $(T_1 + S_1, \ldots, T_n + S_n)$ is NDS.

Remark. Karlin and Rinott (1980) have proven a similar result. They assumed that $\underline{S}$ and $\underline{T}$ satisfy their RR_2 condition and that they have PF_2 marginals, and they showed that then $\underline{S} + \underline{T}$ satisfy some inequalities that are essentially variants of the NUOD and the NLOD inequalities.

For the proof of Theorem 4.4 we need the following lemmas.

Lemma 4.1. Let X and Y be independent rv's and assume that Y has a PF_2 density (probability function). Then

$$[X \mid X + Y = z] \overset{st}{\uparrow} z.$$

This lemma is actually Example 12 of Lehmann (1966).

Lemma 4.2. Let X and Y be independent and assume that X and Y have PF_2 densities (probability functions). Then

$$[(X,Y) \mid X + Y = z] \overset{st}{\uparrow} z. \tag{4.2}$$

Proof. Let $z_1 \leq z_2$. Denote by $U_{(x,y)}$ the upper set $\{(s,t): s > x,\ t > y\}$. First it will be shown that for every (x,y)

$$P\{(X,Y) \in U_{(x,y)} \mid X + Y = z_1\} \leq P\{(X,Y) \in U_{(x,y)} \mid X + Y = z_2\}. \tag{4.3}$$

Let $A_i = \{(s,t): s + t = z_i,\ x < s < z_i - y\}$,

$B_i = \{(s,t): s + t = z_i,\ s \geq z_i - y\}$, $C_i = \{(s,t): s + t = z_i,\ s \leq x\}$,

$i = 1,2$ (see Figure 1). From Lemma 4.1 $[X|X + Y = z] \overset{st}{\uparrow} z$; hence

$$P\{(X,Y) \in C_1|X + Y = z_1\} = P\{X \leq x|X + Y = z_1\}$$

$$\geq P\{X \leq x|X + Y = z_2\} = P\{(X,Y) \in C_2|X + Y = z_2\}.$$

Similarly

$$P\{(X,Y) \in B_1|X + Y = z_1\} \geq P\{(X,Y) \in B_2|X + Y = z_2\}.$$

Thus

$$P\{(X,Y) \in A_1|X + Y = z_1\} \geq P\{(X,Y) \in A_2|X + Y = z_2\}$$

which is (4.3).

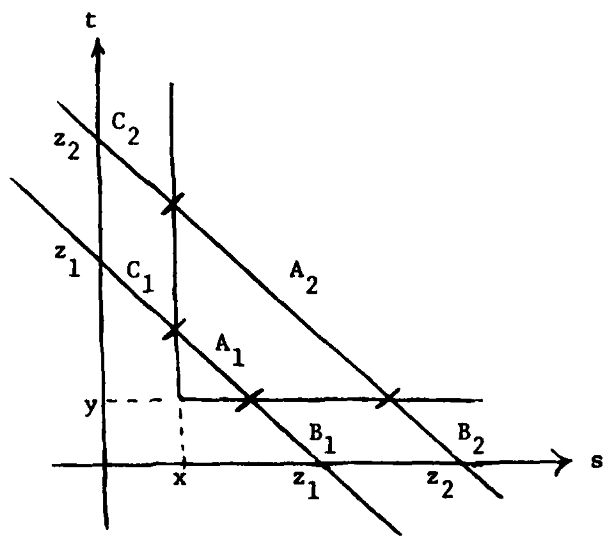


Figure 1.

Now consider upper sets of the form $U = \bigcup_{i=1}^{m} U_{(x_i,y_i)}$, called fundamental upper domains in Block and Savits (1979). Without loss of generality assume that $x_1 \leq x_2 \leq \dots \leq x_m$ and $y_1 \geq y_2 \geq \dots \geq y_m$. Define

$$i_1 = \min\{i: x_i + y_i \leq z_2\}$$

$$j_1 = \min\{i > i_1: x_i + y_i > z_2\}$$

and, by induction

$$i_{k+1} = \min\{i > j_k: x_i + y_i \leq z_2\} \tag{4.4}$$

$$j_{k+1} = \min\{i > i_{k+1}: x_i + y_i > z_2\} \tag{4.5}$$

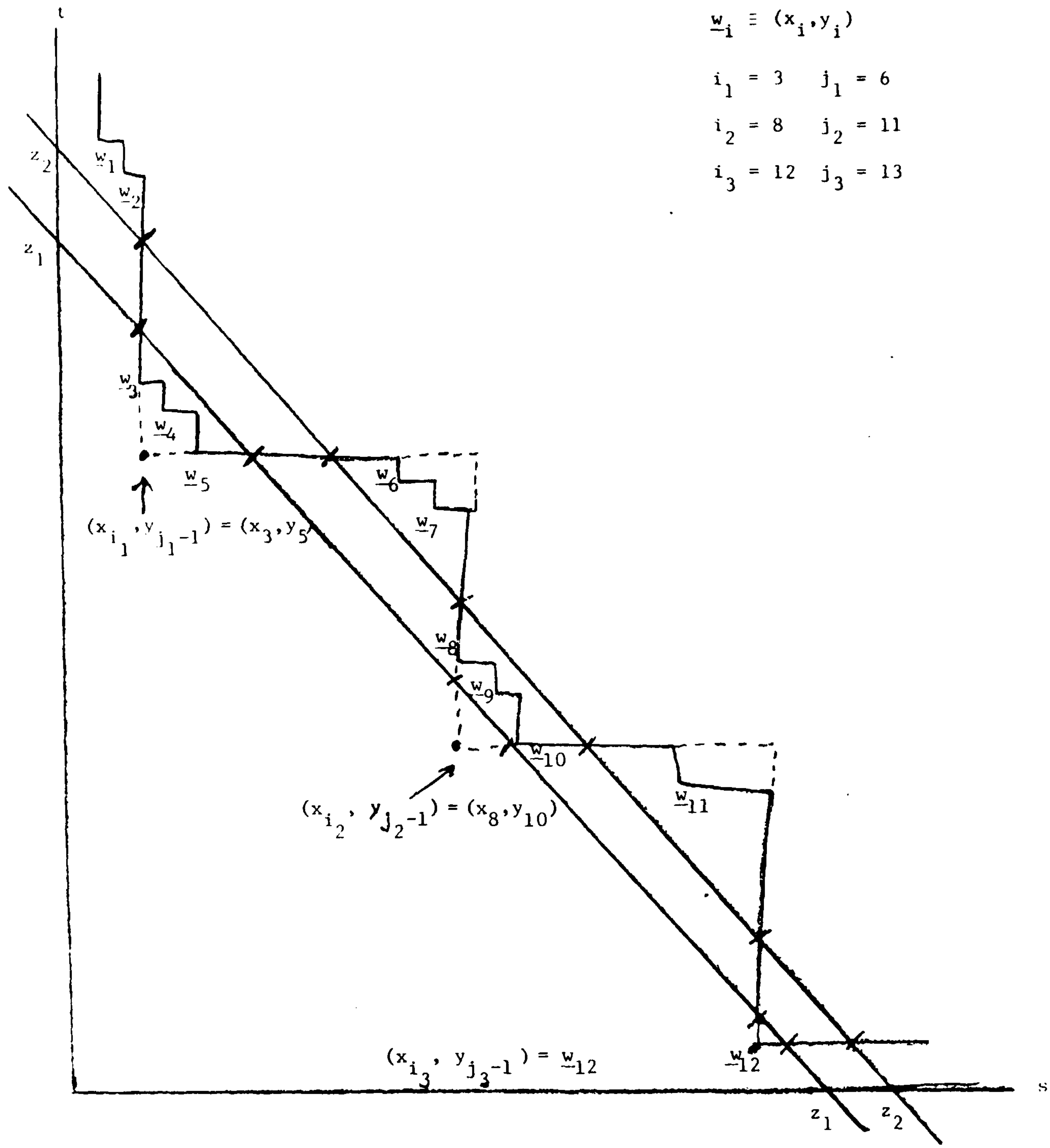
$k = 1,2,\dots,n$ where $n < \infty$ is the largest k such that the set on the right hand side (RHS) of (4.4) or (4.5) is not empty. If there are n i's but only $n-1$ j's define $j_n = m + 1$ (see Figure 2 in which $m = 12$, $n = 3$, $i_1 = 3$, $j_1 = 6$, $i_2 = 8$, $j_2 = 11$, $i_3 = 12$, $j_3 = 13$).

Let $\tilde{U} = \bigcup_{k=1}^{n} U_{(x_{i_k}, y_{j_{k-1}})}$. In the following the first inequality follows from $U \subseteq \tilde{U}$ and the second one from (4.3):

$$P\{(X,Y) \in U | X + Y = z_1\} \leq P\{(X,Y) \in \tilde{U} | X + Y = z_1\}$$

$$\begin{aligned} &= \sum_{k=1}^{n} P\{(X,Y) \in U_{(x_{i_k}, y_{j_k - 1})} | X + Y = z_1\} \\ &\leq \sum_{k=1}^{n} P\{(X,Y) \in U_{(x_{i_k}, y_{j_k - 1})} | X + Y = z_2\} \end{aligned} \tag{4.6}$$

$$= P\{(X,Y) \in \tilde{U} | X + Y = z_2\} = P\{(X,Y) \in U | X + Y = z_2\}.$$



$\underline{w}_i \equiv (x_i, y_i)$
$i_1 = 3 \quad j_1 = 6$
$i_2 = 8 \quad j_2 = 11$
$i_3 = 12 \quad j_3 = 13$
t
s
z_1
z_2
$\underline{w}_1$
$\underline{w}_2$
$\underline{w}_3$
$\underline{w}_4$
$\underline{w}_5$
$\underline{w}_6$
$\underline{w}_7$
$\underline{w}_8$
$\underline{w}_9$
$\underline{w}_{10}$
$\underline{w}_{11}$
$\underline{w}_{12}$
$(x_{i_1}, y_{j_1-1}) = (x_3, y_5)$
$(x_{i_2}, y_{j_2-1}) = (x_8, y_{10})$
$(x_{i_3}, y_{j_3-1}) = \underline{w}_{12}$

Figure 2.

Since open upper sets can be approximated by fundamental upper domains (Block and Savits (1979)) it follows from (4.6) that for every open upper set U

$$P\{(X,Y) \in U \mid X + Y = z_1\} \leq P\{(X,Y) \in U \mid X + Y = z_2\}$$

and the proof of the lemma is complete by (1.1). ||

Lemma 4.3. Let $\underline{X} = (X_1,\ldots,X_m)$ and $\underline{Y} = (Y_1,\ldots,Y_m)$ be independent and assume

(4.7) $$[(X_1,\ldots,X_{m-1}) \mid X_m = x_m] \overset{st}{\downarrow} x_m$$

and

(4.8) $$[(Y_1,\ldots,Y_{m-1}) \mid Y_m = y_m] \overset{st}{\downarrow} y_m.$$

Furthermore, assume that X_m and Y_m have PF_2 densities (probability functions). Then

$$[(X_1 + Y_1,\ldots,X_{m-1} + Y_{m-1}) \mid X_m + Y_m = z_m] \overset{st}{\downarrow} z_m.$$

Proof. Clearly, for any increasing function g,
$E[g(X_1 + Y_1,\ldots,X_{m-1} + Y_{m-1}) \mid X_m + Y_m = z_m] = E[\phi(X_m,Y_m) \mid X_m + Y_m = z_m]$ where $\phi(x_m,y_m) = E[g(X_1 + Y_1,\ldots,X_{m-1} + Y_{m-1}) \mid X_m = x_m,\ Y_m = y_m]$. However, $\phi(x_m,y_m)$ decreases in x_m and in y_m because of (4.7), (4.8) and independence. Thus, by Lemma 4.2, $E[\phi(X_m,Y_m) \mid X_m + Y_m = z_m]$ decreases in z_m. ||

Proof of Theorem 4.4. Let $i \in \{1,\ldots,n-1\}$. Substitute $m = i + 1$ in

Lemma 4.3 to obtain

$$[(T_1 + S_1,\ldots,T_i + S_i)\,|\,T_{i+1} + S_{i+1} = z_{i+1}] \overset{st}{\downarrow} z_{i+1},$$

that is, $\underline{T} + \underline{S}$ is NDS. ||

5. Examples.

5.1. The multinomial df.

Let $(T_1,\ldots,T_n)$ have the joint probability function with parameters $(N,p_1,\ldots,p_n)$,

$$P(T_1 = t_1,\ldots,T_n = t_n) = \frac{N!}{t_1!\ldots t_n!(N-\sum_{i=1}^{n} t_i)!}\,(\prod_{i=1}^{n} p_i^{t_i})$$

$$\times\ (1-\sum_{i=1}^{n} p_i)^{N-\sum_{i=1}^{n} t_i},\quad t_i \geq 0,\quad \sum_{i=1}^{n} t_i \leq N,$$

where $p_i \geq 0$ $(i = 1,\ldots,n)$ and $0 < \sum_{i=1}^{n} p_i < 1$.

The *multinomial* df is the conditional df of independent Poisson rv's given their sum. Thus, by Theorem 4.3 the multinomial df is RR_2 in pairs and hence it is also CDS, NUOD and NLOD. By Remark (iii) the joint probability function of $(T_1,\ldots,T_n)$ is RR_2 in pairs. By the discussion after Theorem 4.3 the multinomial df satisfies the RR_2 condition of Karlin and Rinott (1980).

To show that $(T_1,\ldots,T_n)$ is NDS it is enough to show that for $0 \leq t_n \leq N-1$,

(5.1.1) $$[(T_1,\ldots,T_{n-1})\,|\,T_n = t_n] \overset{st}{\geq} [(T_1,\ldots,T_{n-1})\,|\,T_n = t_n+1]$$

because then, by symmetry one has (2.4'). The left hand side (LHS) of (5.1.1) has an $(n-1)$-dimensional multinomial df with parameters $(N-t_n, q_1, \ldots, q_{n-1})$ and the right hand side (RHS) of (5.1.1) has the same df with parameters $(N-t_n-1, q_1, \ldots, q_{n-1})$ where $q_i = p_i/(1-p_n)$. Thus,

$$\text{(5.1.2)} \qquad \text{LHS (5.1.1)} \overset{st}{=} \text{RHS (5.1.1)} + (S_1, \ldots, S_{n-1})$$

where $(S_1, \ldots, S_{n-1})$ has a multinomial df with parameters $(1, q_1, \ldots q_{n-1})$ and the sum on the RHS of (5.1.2) is of independent random vectors. Inequality (5.1.1) follows now from (5.1.2) and (1.3).

Let $(X_1, \ldots, X_n)$ be the sum of independent n-dimensional, not necessarily identically distributed, multinomial random vectors, that is, let $(X_1, \ldots, X_n)$ have the probability generating function

$$\prod_{\ell=1}^{m} (p_{1\ell}u_1 + \ldots + p_{n\ell}u_n)^{N_\ell}$$

where $p_{i\ell} \geq 0$, $(i = 1, \ldots, n)$, $0 < \sum_{i=1}^{n} p_{i\ell} < 1$ and $N_\ell \geq 1$ $(\ell = 1, 2, \ldots, m)$. The univariate marginal df's of a multinomial df is a binomial df. Since the binomial df has a PF_2 probability function it follows from Theorem 4.4 that $(X_1, \ldots, X_n)$ is NDS. Hence $(X_1, \ldots, X_n)$ is also NUOD and NLOD, that is,

$$P(X_1 > x_1, \ldots, X_n > x_n) \leq \prod_{i=1}^{n} P(X_i > x_i)$$

and

$$P(X_1 \leq x_1, \ldots, X_n \leq x_n) \leq \prod_{i=1}^{n} P(X_i \leq x_i).$$

These inequalities are stated unproven in Mallows (1968). Compare also Lehmann (1966), pp. 1143, 1144 and 1151.

5.2. Multivariate normal.

Let $\underset{\sim}{T} = (T_1,\dots,T_n)$ be a multivariate symmetric normal random vector with $\mathrm{Corr}(T_i,T_j) = \rho \le 0$, $1 \le i < j \le n$. Then $\rho \ge -(n-1)^{-1}$. We will show that $\underset{\sim}{T}$ is RR_2 in pairs.

Using Theorem 4.1 assume, without loss of generality, that $ET_i = 0$ and $\mathrm{Var}(T_i) = 1$, $i = 1,\dots,n$. Let $Y_1,\dots,Y_n$ be independent identically distributed normal rv's such that $EY_i = 0$ and $\mathrm{Var}(Y_i) = 1-\rho$ $(i = 1,\dots,n)$ and let Y_0 be an independent normal rv with $EY_0 = 0$ and $\mathrm{Var}(Y_0) = (-\rho)^{-1}(1-\rho)(1+(n-1)\rho)$. Then

$$(T_1,\dots,T_n) \overset{st}{=} [(Y_1,\dots,Y_n) \mid Y_0 + Y_1 + \dots + Y_n = 0].$$

Since any normal density is PF_2 it follows, from Theorem 4.3, that $\underset{\sim}{T}$ is RR_2 in pairs.

In fact we can obtain a stronger result. If the correlation matrix of $\underset{\sim}{T}$ is of the form

$$(5.2.1) \qquad \begin{pmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_n \end{pmatrix} - \begin{pmatrix} (r_1-1)^{1/2} \\ \vdots \\ (r_n-1)^{1/2} \end{pmatrix} ((r_1-1)^{1/2},\dots,(r_n-1)^{1/2}) \; .$$

where $r_i \ge 1$, $i = 1,\dots,n$ and $\sum_{i=1}^{n} r_i^{-1} \ge n-1$ then $\underset{\sim}{T}$ is RR_2 in pairs.

To show it note that every matrix of the form (5.2.1) can be the correlation matrix of a multinomial random vector, $\underset{\sim}{X} = (X_1,\dots,X_n)$, say. Let $\underset{\sim}{X}^{(\ell)}$, $\ell = 1,2,\dots,$ be a sequence of independent random vectors distributed as $\underset{\sim}{X}$. Clearly $\underset{\sim}{Y}^{(m)} = \sum_{\ell=1}^{m} \underset{\sim}{X}^{(\ell)}$ is a multinomial random vector with correlation matrix (5.2.1). Normalizing $\underset{\sim}{Y}^{(m)}$ such that it has zero means and unit variances, it converges in distribution, by the multivariate central limit theorem, to a multivariate normal random vector with correlation matrix (5.2.1). By Remark (vii) the limit in distribution of RR_2 in pairs random vectors is RR_2 in pairs. The assertion in the preceeding paragraph now follows. The previous result that deals with the symmetric multivariate normal df with negative correlations ρ is obtained by taking $r = 1 - \rho$ in (5.2.1).

Every multivariate normal random vector $\underline{T}$ with nonpositive correlations is NDS. It is easy to verify (2.4') directly using (1.2). Thus $\underline{T}$ satisfies (2.3) and (2.3').

5.3. Multivariate hypergeometric.

Let $(T_1,\ldots,T_n)$ have the probability function

$$P(T = t_1,\ldots,T_n = t_n) = \binom{M}{N}^{-1} \left[\prod_{i=1}^{n} \binom{M_i}{t_i}\right] \binom{M - \sum_{i=1}^{n} M_i}{M - \sum_{i=1}^{n} t_i}, \quad t_i \geq 0, \quad \sum_{i=1}^{n} t_i \leq N,$$

with positive integer-valued parameter vector $(N, M_1,\ldots,M_n, M)$ [see Johnson and Kotz (1969)].

The multivariate hypergeometric df is the conditional df of independent binomial rv's given their sum. Thus, by Theorem 4.3 the hypergeometric df is RR_2 in pairs and hence it is also CDS, NUOD and NLOD. By Remark (iii) the joint probability function of $(T_1,\ldots,T_n)$ is RR_2 in pairs. A special case of this fact was observed by Lehmann (1966), p. 1144. See also Ebrahimi and Ghosh (1980). By the discussion after Theorem 4.3 it follows that this df satisfies the RR_2 condition of Karlin and Rinott (1980).

We will show now that $\underline{T}$ is NDS by proving that, for $0 \le t_n \le N-1$,

$$[(T_1,\dots,T_{n-1})|T_n = t_n] \overset{st}{\ge} [(T_1,\dots,T_{n-1})|T_n = t_n + 1]. \tag{5.3.1}$$

Then, by symmetry, (2.4') follows. Since the random vectors on each side of (5.3.1) have multivariate hypergeometric df we actually have to show that

$$\underline{U} \overset{st}{\ge} \underline{V} \tag{5.3.2}$$

where $\underline{U} = (U_1,\dots,U_{n-1})$ has hypergeometric df with parameters $(\tilde{N} + 1, M_1,\dots,M_{n-1}, \tilde{M} + 1)$ and $\underline{V} = (V_1,\dots,V_{n-1})$ has the same df with parameters $(\tilde{N}, M_1,\dots,M_{n-1}, \tilde{M})$ where $\tilde{N} = N - t_n - 1$ and $\tilde{M} = M - t_n - 1$.

Denote $m = n - 1$.

Thinking about U_i as the number of individuals in the sample in Category i $(i = 1,\dots,m)$ it is easy to see, by conditioning on the category of the first individual chosen, that $(U_1,\dots,U_m) \overset{st}{=} (W_1^{(\ell)},\dots,W_{\ell-1}^{(\ell)}, W_\ell^{(\ell)} + 1, W_{\ell+1}^{(\ell)},\dots,W_m^{(\ell)})$ if the first individual is in category ℓ, $\ell = 1,2,\dots,m$, and that $(U_1,\dots,U_m) \overset{st}{=} (V_1,\dots,V_m)$ if the first individual is in neither of categories $1,2,\dots,m$, where $(W_1^{(\ell)},\dots,W_m^{(\ell)})$ has a multivariate hypergeometric df with parameters $(\tilde{N}, M_1,\dots,M_{\ell-1}, M_\ell-1, M_{\ell+1},\dots,M_m, \tilde{M})$. Thus for proving (5.3.2) it suffices to show that

$$(W_1^{(\ell)},\dots,W_{\ell-1}^{(\ell)}, W_\ell^{(\ell)}+1, W_{\ell+1}^{(\ell)},\dots,W_m^{(\ell)}) \overset{st}{\ge} (V_1,\dots,V_m),\ \ell = 1,\dots,m. \tag{5.3.3}$$

We will prove (5.3.3) when $\ell = 1$; the proof for the other ℓ's is similar. Omit the superscript 1 and consider $(W_1,\dots,W_m) = (W_1^{(1)},\dots,W_m^{(1)})$

which has a hypergeometric distribution with parameters $(\tilde{N}, M_1 - 1, M_2, \ldots, M_{m-1}, M_m, \tilde{M})$.

Write

$$P(V_1 = v_1, \ldots, V_m = v_m) = \binom{\tilde{M}}{M_1, \ldots, M_m, \tilde{M} - \sum_{i=1}^{m} M_i}^{-1}$$

$$\times \binom{\tilde{N}}{v_1, \ldots, v_m, \tilde{N} - \sum_{i=1}^{m} v_i} \binom{\tilde{M} - \tilde{N}}{M_1 - v_1, \ldots, M_m - v_m, \tilde{M} - \tilde{N} - \sum_{i=1}^{m} M_i + \sum_{i=1}^{m} v_i}$$

Then, it is easily seen, that the rv's $V_1, \ldots, V_m$ can have the following interpretation: First divide a population of size $\tilde{M}$ into a group of size $\tilde{N}$ and another group of size $\tilde{M} - \tilde{N}$. Next, choose at random M_1 individuals and let V_1 be the number of them in the first group (of size $\tilde{N}$), then choose at random M_2 individuals of the remaining $\tilde{M} - M_1$ individuals and let V_2 be the number of them in the first group. Continuing this way, finally choose M_m individuals of the remaining $\tilde{M} - M_1 - \ldots - M_{m-1}$ individuals and let V_m be the number of them in the first group.

The rv's $W_1, \ldots, W_m$ may have a similar interpretation.

With these interpretations it is easily seen that if the first individual chosen in the '$\underline{V}$ experiment' is not in the first group then $\underline{V} \overset{st}{=} \underline{W} \overset{st}{<} (W_1 + 1, W_2, \ldots, W_m)$, where the inequality follows from (1.2). Otherwise $(V_1, \ldots, V_m) \overset{st}{=} (W_1 + 1, W_2, \ldots, W_m)$. Thus, unconditionally $\underline{V} \overset{st}{<} (W_1 + 1, \ldots, W_m)$, which proves (4.3.3) when $\ell = 1$.

5.4. The Dirichlet distribution.

Let $\underline{T} = (T_1,\dots,T_n)$ have the density

$$f(t_1,\dots,t_n) = \frac{\Gamma(\sum_{j=0}^{n}\theta_j)}{\prod_{j=0}^{n}\Gamma(\theta_j)}\,(1-\sum_{j=1}^{n}t_j)^{\theta_0-1}\prod_{j=1}^{n}t_j^{\theta_j-1},\quad t_j \ge 0,\quad \sum_{j=1}^{n}t_j \le 1.$$

where the parameter vector $(\theta_0,\theta_1,\dots,\theta_n)$ satisfies $\theta_j \ge 1$, $j = 0,1,\dots,n$.

The Dirichlet df is the conditional df of independent gamma rv's given their sum. Thus, by Theorem 4.3 the Dirichlet df is RR_2 in pairs and hence it is also CDS, NUOD and NLOD. By Remark (iii) f is RR_2 in pairs. Special case of this fact is Example 10 (iii) of Lehmann (1966). See also Ebrahimi and Ghosh (1980). By the discussion after Theorem 4.3 it follows that this df satisfies the RR_2 condition of Karlin and Rinott (1980).

To prove that $\underline{T}$ is NDS it is enough to show that for $t_n \le t_n'$

(5.4.1) $$[(T_1,\dots,T_{n-1})|T_n = t_n] \overset{st}{\ge} [(T_1,\dots,T_{n-1})|T_n = t_n'],$$

then (2.4') follows.

It is easily seen that

$$(1-t_n)^{-1}[(T_1,\dots,T_{n-1})|T_n = t_n] \overset{st}{=} (1-t_n')^{-1}[(T_1,\dots,T_{n-1})|T_n = t_n'].$$

Since $1 - t_n \ge 1 - t_n'$ (5.4.1) follows from (1.4).

5.5. Dirichlet compound multinomial df.

Let $\underline{T} = (T_1,\dots,T_n)$ have the probability function

$$P(T_1 = t_1,\dots,T_n = t_n) = \frac{N!\,\Gamma(\sum_{j=0}^{n}\theta_j)}{\Gamma(N + \sum_{j=0}^{n}\theta_j)} \prod_{j=1}^{n} \frac{\Gamma(t_j+\theta_j)}{t_j!\,\Gamma(\theta_j)}$$

$$\times \frac{\Gamma(N - \sum_{j=1}^{n} t_j + \theta_0)}{(N - \sum_{j=1}^{n} t_j)!\,\Gamma(\theta_0)}, \quad t_j \geq 0, \quad \sum_{j=1}^{n} t_j \leq N,$$

where N is a positive integer and $\theta_j \geq 1$, $j = 0,1,\dots,n$ [see Johnson and Kotz (1969)].

The Dirichlet compound multinomial df is the conditional df of independent Pascal (negative binomial) rv's given their sum. Thus, by Theorem 4.3 this df is RR_2 in pairs and hence it is also CDS, NUOD and NLOD, and by the remark after Theorem 4.3 it satisfies the RR_2 condition of Karlin and Rinott (1980).

To show that $\underline{T}$ is NDS it suffices to prove that

$$[(T_1,\dots,T_{n-1})|T_n = t_n] \overset{st}{\downarrow} t_n, \tag{5.5.1}$$

then (2.4') follows by symmetry. But the random vector in (5.5.1) has the Dirichlet compound multinomial df with parameters $N - t_n$, $\theta_0,\dots,\theta_{n-1}$. Using this fact it is easy to verify (5.5.1).

5.6. Df's supported by negatively tilted surfaces.

In this subsection we illustrate two simple examples of such df's in $\mathbb{R}^3$. The general idea then should be clear and will not be detailed here.

Let (T_1,T_2,T_3) have the uniform df on the surface $\{(t_1,t_2,t_3) \geq (0,0,0): t_1 + t_2 + t_3 = 1\}$. Then $[(T_1,T_2)|T_3 = t_3]$ has a uniform df on the segment $\{(t_1,t_2) \geq (0,0): t_1 + t_2 = 1 - t_3\}$ which is clearly stochastically decreasing in t_3. By symmetry (2.4') holds, hence (T_1,T_2,T_3) is NDS. Clearly (T_1,T_2,T_3) is also CDS.

Similarly, if (T_1,T_2,T_3) have the uniform df on the surface of the unit sphere intersected with the positive (or the negative) orthant then (T_1,T_2,T_3) is NDS and CDS.

5.7. NUOD df's in reliability theory.

Buchanan and Singpurwalla (1977) considered the class of nonnegative multivariate new better than used (NBU) df's that satisfy, for all $\underline{s} > \underline{0}$, $\underline{t} > \underline{0}$,

$$(5.7.1) \qquad P\{T_1 > s_1 + t_1,\dots,T_n > s_n + t_n\} \leq P(T_1 > s_1,\dots,T_n > s_n) \times P(T_1 > t_1,\dots,T_n > t_n).$$

It is well known that, when $n = 2$, (5.7.1) implies that (T_1,T_2) are negatively quadrant dependent, that is

$$(5.7.2) \qquad P(T_1 > t_1, T_2 > t_2) \leq P(T_1 > t_1)\, P(T_2 > t_2).$$

We will show, by induction, that in general (5.7.1) implies that $(T_1,\dots,T_n)$ is NUOD. From (5.7.2) we know it is true when $n = 2$. Assume that

$$(5.7.3) \qquad P(T_1 > t_1,\dots,T_{n-1} > t_{n-1}) \leq \prod_{i=1}^{n-1} P(T_i > t_i).$$

Substitute $s_1 = s_2 = \ldots = s_{n-1} = t_n = 0$ in (5.7.1) and use (5.7.3) to obtain

$$P(T_1 > t_1,\ldots,T_{n-1} > t_{n-1}, T_n > s_n) \le P(T_n > s_n)\, P(T_1 > t_1,\ldots,T_{n-1} > t_{n-1})$$

$$\le \left[\prod_{i=1}^{n-1} P(T_i > t_i)\right] P(T_n > s_n),$$

that is, T is NUOD.

References

[1] Ahmed, A.H.N., Langberg, N.A., Leon, R.V. and Proschan, F. (1978). Two concepts of positive dependence, with applications in multivariate analysis. Technical Report, Department of Statistics, The Florida State University.

[2] Arjas, E. and Lehtonen, T. (1978). Approximating many server queues by means of single server queues. Mathematics of Operations Research, 3, 205-223.

[3] Barlow, R.E. and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing: Probability Models. Holt, Rinehart and Winston, New York.

[4] Block, H.W. and Savits, T.H. (1979). Multivariate IFRA distribution. Ann. Prob., to appear.

[5] Brindley, E.C. and Thompson, W.A. (1972). Dependence and aging aspects of multivariate survival. J. Amer. Statist. Assoc., 67, 822-830.

[6] Buchanan, W.B. and Singpurwalla, N.D. (1977). Some stochastic characterizations of multivariate survival. The Theory and Applications of Reliability (Eds.: C.P. Tsokos and I.N. Shimi), 329-348.

[7] Dykstra, R.L., Hewett, J.E. and Thompson, W.A. (1973). Events which are almost independent. Ann. Statist., 1, 674-681.

[8] Ebrahimi, N. and Ghosh, M. (1980). Multivariate negative dependence. Technical Report, Department of Statistics, Iowa State University, Ames, Iowa.

[9] Johnson, N.L. and Kotz, S. (1969). Distribution in Statistics, Discrete Distributions. Wiley, New York.

[10] Kamae, T., Krengel, U. and O'Brien, G.L. (1977). Stochastic inequalities on partially ordered spaces. Ann. Prob., 5, 899-912.

[11] Karlin, S. (1968). Total Positivity. Stanford University Press, Stanford.

[12] Karlin, S. and Rinott, Y. (1980). Classes of orderings of measures and related correlation inequalities. II. multivariate reverse rule distributions. Technical Report, Department of Mathematics, Stanford University.

[13] Kemperman, J.H.B. (1977). On the FKG-inequality measures on a partially ordered space. Proc. Kon. Ned. Akad. Wet., Amsterdam, Series A, Vol. 80 (4), 313-331.

[14] Lehmann, E.L. (1966). Some concepts of dependence. Ann. Math. Statist. 43, 1137-1153.

[15] Mallows, C.L. (1968). An inequality involving multinomial probabilities. Biometrika, 55, 422-424.

[16] Meyer, P.A. (1966). <u>Probability and Potentials</u>. Blaisdell, Waltham, Mass..

[17] Shaked, M. (1977). A family of concepts of positive dependence for bivariate distributions. <u>J. Amer. Statist. Assoc.</u>, <u>72</u>, 642-650.

DATE

ILMED

-8